Impact of sialyltransferase ST6GAL1 overexpression on different colon cancer cell types

Giulia Venturi[1]*, Inês Gomes Ferreira[1]*, Michela Pucci[1], Manuela Ferracin[1], Nadia Malagolini[1], Mariella Chiricolo[1] Fabio Dall'Olio[1]**

[1]Department of Experimental, Diagnostic and Specialty Medicine, General Pathology Building, University of Bologna, Via S. Giacomo 14, 40126, Bologna, Italy

*** Equal contribution**

**** Correspondence:**
Fabio Dall'Olio
fabio.dallolio@unibo.it

Running title: Impact of ST6GAL1 in different colon cancer cell types

Keywords: colorectal cancer/glycosylation/sialylation/sialyltransferases/transcriptomic analysis

Supplementary data:

Figure S I

Figure S II

Table S I

Table S II

Table S III

Abstract

Cancer-associated glycan structures can be both tumor markers and engines of disease progression. The structure Siaα2,6Galβ1,4GlcNAc (Sia6LacNAc), synthesized by sialyltransferase *ST6GAL1*, is a cancer-associated glycan. Although *ST6GAL1*/Sia6LacNAc are often overexpressed in colorectal cancer (CRC), their biological and clinical significance remains unclear. To get insights into the clinical relevance of *ST6GAL1* expression in CRC, we interrogated The Cancer Genome Atlas (TCGA) with mRNA expression data of hundreds of clinically characterized CRC and normal samples. We found an association of low *ST6GAL1* expression with microsatellite instability, *BRAF* mutations and mucinous phenotype but not with stage, response to therapy and survival. To investigate the impact of *ST6GAL1* expression in experimental systems, we analyzed the transcriptome and the phenotype of the CRC cell lines SW948 and SW48 after retroviral transduction with *ST6GAL1* cDNA. The two cell lines display the two main pathways of CRC transformation: chromosomal instability and microsatellite instability, respectively. Constitutive *ST6GAL1* expression induced much deeper transcriptomic changes in SW948 than in SW48 and affected different genes in the two cell lines. *ST6GAL1* expression affected differentially the tyrosine phosphorylation induced by hepatocyte growth factor, the ability to grow in soft agar, to heal a scratch wound and to invade Matrigel in the two cell lines. These results indicate that the altered expression of a cancer-associated glycosyltransferase impacts the gene expression profile, as well as the phenotype, although in a cancer subtype-specific manner.

Introduction

Glycosylation is a fundamental post-translational protein modification, which undergoes profound changes in cancer (Dall'Olio et al. 2012; Pinho et al. 2015). Altered sialylation is one of the most relevant cancer-associated glycosylation changes and is mainly due to the altered expression of sialyltransferases (Dall'Olio et al. 2014). Sialyltransferase ST6GAL1, product of the *ST6GAL1* gene, is the only enzyme able to add sialic acid on the subterminal galactose residues of lactosaminic chains (Galβ1,4GlcNAc) of glycoproteins (Weinstein et al. 1987), resulting in the biosynthesis of Sia6LacNAc (Siaα2,6Galβ1,4GlcNAc) (Supplementary Fig. S IA). Sia6LacNAc is frequently found at the non-reducing termini of N-linked chains of glycoproteins and can be detected by the lectin from *Sambucus nigra* (SNA) (Shibuya et al. 1987). The increased expression of ST6GAL1 in colorectal cancer (CRC) tissues was formerly reported by our laboratory (Dall'Olio et al. 1989) and later confirmed by others (Lise et al. 2000; Petretti et al. 2000; Seales et al. 2005). Increased expression of ST6GAL1 and/or of its cognate Sia6LacNAc structure has been observed in a variety of other malignancies (reviewed in (Dall'Olio 2000; Dall'Olio et al. 2001; Dall'Olio et al. 2014)) of epithelial (Fukushima et al. 1998; Jun et al. 2012; Schultz et al. 2016; Shah et al. 2008) and non-epithelial origin (Mondal et al. 2010; Skacel et al. 1991). The biological and clinical implications of ST6GAL1 up-regulation have been the focus of numerous experimental and clinical studies. Some clinical CRC investigations report a positive relationship between high ST6GAL1/ Sia6LacNAc and increased malignancy (Gessner et al. 1993; Lise et al. 2000; Vierbuchen et al. 1995), although some show the opposite (Yamashita et al. 1995; Zhang et al. 2017). Experimental studies on CRC cell lines genetically manipulated to up- or downregulate ST6GAL1, indicate that this enzyme positively affects resistance to apoptosis (Swindall et al. 2011; Zhuo et al. 2008), drug resistance (Chang et al. 2018; Cui et al. 2018), radiation resistance (Lee et al. 2008; Lee et al. 2010a; Lee et al. 2010b) and stemness (Cui et al. 2018; Swindall et al. 2013). However, *in vivo* growth (Chiricolo et al. 2006; Park et al. 2012) and metastasis (Jung et al. 2016) were found to be higher in the absence of ST6GAL1. In other cancers, ST6GAL1 was found to support an invasive phenotype according to several studies (Britain et al. 2017; Britain et al. 2018; Chakraborty et al. 2018; Han et al. 2018; Jones et al. 2018; Meng et al. 2015; Schultz et al. 2016; Wang et al. 2015; Wei et al. 2016), but not by others (Antony et al. 2014; Dawson et al. 2004; Yamamoto et al. 1997; Yamamoto et al. 2001; Yen et al. 2015).This study was undertaken to clarify the role of ST6GAL1 in CRC both in the clinic and in experimental systems. To this aim, we interrogated The Cancer Genome Atlas (TCGA) data portal, a public database containing the gene expression data of hundreds of clinically and molecularly characterized

CRC fresh tumors. Second, we investigated the effect of ST6GAL1 expression on the phenotype and

gene expression of two CRC cells lines, which display the two main pathways of CRC transformation: chromosomal instability (CIN) and microsatellite instability (MSI).

Results

Survey of transcriptomic databases

The analysis of transcriptomic data of hundreds of CRC and normal specimens from TCGA database allowed to investigate the relationship between *ST6GAL1* gene expression and clinical parameters. As shown in Fig. 1A and B, *ST6GAL1* mRNA expression is variable among CRC specimens (A) but quite uniform in normal tissues (B). No relationship existed between *ST6GAL1* expression and clinical stage (Fig. 1C), while a highly significant association was found between low *ST6GAL1* expression and high microsatellite instability (MSI-h) (Fig. 1D). We did not observe an association between *ST6GAL1* expression and *APC* (Fig. 1E), *TP53* (Fig. 1F) and *KRAS* (Fig. 1G) mutations, but we found a highly significant association between low *ST6GAL1* expression and *BRAF* mutation (Fig. 1H). Low *ST6GAL1* was associated also with a mucinous phenotype (Fig. 1I), while no relationship existed between *ST6GAL1* expression and response to therapy (Fig. 1J) or overall survival (Fig. 1K). To strength the relationship between *ST6GAL1* expression and microsatellite status, a different cohort of 23 MSS and 16 MSI-h CRC cases was investigated by a different microarray technology (Ferracin et al. 2008). As reported in Fig. S I, the lower ST6GALI expression expressed by MSI-h cases is fully consistent with that observed in the TCGA cohort. Collectively, these data suggest an association of *ST6GAL1* expression with specific CRC features, but not with general progression. In search of a gene expression signature associated with high or low *ST6GAL1* expression, two cohorts including the patients in the 15% upper and 15% lower percentiles of *ST6GAL1* mRNA level were compared. As shown in Table I, the two cohorts displayed completely different gene expression signatures. In fact, samples from the upper percentile showed up-regulation of a few genes, including two CEA-related genes (*CEACAM5* and *CEACAM6*) and *CD24*, whose products share a common involvement in cell adhesion. On the contrary, low *ST6GAL1* patients displayed up-regulation of a number of genes involved in several cellular functions, including carbohydrate metabolism, regeneration and repair, protein synthesis, mucin production and cytoskeleton organization. In particular, three members of the cytokeratin family (*KRT8, KRT18* and *KRT19*) displayed up-regulation in low ST6GAL1 expressers. The high expression of genes related with mucin production in low ST6GAL1 patients is in line with the observation that mucinous cancers display low ST6GAL1 expression (Fig. 1I).

The impact of ST6GAL1 overexpression on the transcriptome is strongly cell-type specific

To establish whether and how the overexpression of *ST6GAL1* affects colon cancer cells, we analyzed the global gene expression profile of SW48 and SW948 colon cancer cell lines, permanently transduced with human *ST6GAL1* cDNA (SW948 ST and SW48 ST) (Malagolini et al. 2009) and their respective negative controls transduced with an empty vector (SW948 NC and SW48 NC). As shown in Supplementary Fig. S II, the negative controls of both cell lines lacked endogenous *ST6GAL1* expression, as previously reported for their wild type counterparts (Dall'Olio et al. 1995). As a result of *ST6GAL1* transduction, both SW948 ST and SW48 ST expressed the *ST6GAL1* mRNA, protein and enzymatic activity and Sia6LacNAc on the cell membrane, as revealed by fluorescent SNA. SNA-lectin blot analysis (Fig. S II F) revealed that multiple glycoproteins of different molecular weights are over α2,6-sialylated in ST6GAL1-transduced SW48 cells. However, in SW948 ST cells the level of over α2,6-sialylation appears to be lower. These two cell lines were chosen because they lack endogenous *ST6GAL1* expression and display the two main pathways of colon cancer transformation. In fact, SW48 cells exhibit microsatellite instability, while SW948 cells display chromosomal instability. The heat map graph shown in Fig. 2 revealed that in both cell lines groups of genes were up- or down-regulated by *ST6GAL1* expression. However, the number of *ST6GAL1*-modulated genes was higher in SW948 than in SW48 cells. Supplementary Table S I shows pathways significantly enriched upon ST6GAL1 overexpression, according to Metacore Pathway Enrichment Analysis. We reported the pathways likely to be relevant in CRC in Table II. In SW948 cells, these include WNT signaling, TGF-β-induced EMT and transcription of HIF-1 target genes. In SW48 cells, the number of putatively affected pathways was much lower, including the metabolism of steroid hormones, the ephrin receptors, and cytoskeleton remodeling. A more in-deep analysis of the function of the *ST6GAL1*-modulated genes in SW948 and SW48 cells was performed through a literature search (Supplementary Tables S II and S III) and their putative tumor promoting or tumor restraining role was marked with a red or green label, respectively. In SW948 ST cells, genes with a recognized CRC-promoting role (*MYEOV, BST2, TGFB2, KLF13, KLF8, XIST, SCGB2A1, NRP2*) and genes with a tumor restraining role in CRC (*ZC3H13, STAT4, EDN2, CTF1*, *WNT4*) displayed up-regulation. Numerous genes with a recognized role in other malignancies displayed up- or down-regulation in SW948 ST cells. In SW48 ST cells, some CRC-promoting genes displayed down-regulation (*TMEM41A, INHHB, INPP4B, EPHA3, ASB4, NTRK2, MAP1B*), while the CRC tumor-restraining genes *PCDH1* and *MUC2* displayed up-regulation. Consistent with TCGA data, the expression of three members of the cytokeratin family (*KRT14*, *KRT16*, *KRT17*) was reduced in *ST6GAL1* expressing SW48 cells. However, this effect, which suggests a possible inhibitory effect of ST6GAL1 on cytokeratin expression, was not observed in SW948 cells, nor on the genes encoding non-cytoskeletal keratins. In SW948 cells, the sum of putative tumor-promoting changes (up-

regulation of red-marked genes + down-regulation of green-marked genes) was similar with that of putative tumor-restraining changes (down-regulation of red-marked genes + up-regulation of green-marked genes) (21 vs.19). On the contrary, in SW48 cells we counted 6 putative tumor-promoting changes vs. 24 putative tumor-restraining changes. Unexpectedly, none of the genes modulated in one cell line showed parallel modulation in the other. Altogether, these data indicate that the gene expression changes induced by *ST6GAL1* overexpression are strong but cell line specific.

HGF-induced tyrosine phosphorylation is differentially modulated by ST6GAL1 in CRC cell lines.

A way through which a glycosyltransferase could influence gene expression, is the modulation of the cell signaling triggered by growth factors. Owing to the well-known role of HGF in stimulating cell motility and migration through FAK phosphorylation (Lai et al. 2000), as a first sight into the mechanisms downstream of ST6GAL1 expression, we investigated the phosphorylation of tyrosine residues and in particular Tyr397 of FAK by Western blotting in the ST6GAL1-transduced and their negative control cell lines after HGF stimulation (Fig. 3). In unstimulated cells, Tyr397 phosphorylation of FAK was significantly higher in SW948 ST than in SW948 NC (Fig. 3A). While in SW948 ST cells pFAK decreased progressively during HGF treatment, in SW948 NC it rose gradually during treatment time. A similar trend was observed on the most abundant Tyr phosphoproteins of 125, 60 and 25 kDa Fig. 3A). On the other hand, treatment of SW48 NC and ST with HGF failed to induce significant changes of Tyr397 phosphorylation of FAK and of the 125, 60 and 25 kDa Tyr-phosphoproteins (Fig. 3B). These results indicate that ST6GAL1 expression strongly affects HGF-induced tyrosine phosphorylation and consequently HGF signaling in SW948 cells but not in SW48 cells.

The phenotype of colon cancer cells is differentially modulated by *ST6GAL1* overexpression.

To investigate the effect of *ST6GAL1* expression on the malignant phenotype, SW48 and SW948 cells transduced with the *ST6GAL1* cDNA and their respective negative controls were assayed for key phenotypic features associated with malignancy. The ability to grow in a semi-solid medium, such as the soft agar (Fig. 4A), was differentially modulated by ST6GAL1 in SW948 and SW48. In fact, it resulted in the formation of fewer clones in the cell line SW948 and in a slightly increased number of clones in the cell line SW48. The time required to heal a scratch wound in a cell culture (Fig. 4B) was shorter for SW948 ST than for SW948 NC. This was largely due to the fact that SW948 ST cells tend to proliferate first in monolayer and only successively as a multilayer, while SW948 NC cells show a stronger tendency to multi-layer growth. On the contrary, SW48 NC cells closed the wound faster than SW48 ST cells. The ability to invade the extracellular matrix (Matrigel) was the 6

same in SW948 NC and ST, while it was strongly reduced in SW48 ST, compared with SW48 NC (Fig. 4C).

ST6GAL1 overexpression and stemness.

To investigate the relationship between *ST6GAL1* expression and stemness, we analyzed the four cells lines for the expression of known markers of stemness: aldehyde dehydrogenase (ALDH) and CD133. The first is reported to be a stem cell and a cancer-initiating cell marker in many tissues, including colon (Volonte et al. 2014) while the second is a surface protein associated with stemness in various cancers, including CRC (Ricci-Vitiani et al. 2007). In a typical experiment (Fig. 5A), cells were incubated with the ALDH substrate ALDEFLUOR, either in the presence or in the absence of DEAB (a specific ALDH inhibitor) for negative control. While the percentage of ALDH-positive cells was about 70% in the SW948 cell type, it was about-30-40% in the SW48 cell type. However, little and non-significant changes attributable to ST6GAL1 expression were observed. Flow cytometric analysis of CD133-labelled cells (Fig. 5B) revealed a slightly higher percentage of CD133+ and CD133+ high cells in the SW948 ST cell line, but no differences between SW48 NC and SW48 ST.

Discussion

To clarify the role of *ST6GAL1* into the clinic, we correlated the *ST6GAL1* mRNA expression with patients' parameters using the TCGA database. The low level of *ST6GAL1* mRNA we observed in high MSI patients in two different cohorts is consistent with a previous observation reporting that high SNA reactivity was correlated with MSS phenotype (Gebert et al. 2012). The association between low *ST6GAL1* and *BRAF* mutation is consistent with the fact that *BRAF* is frequently mutated in MSI CRC patients. Although correlations have been reported in CRC between *ST6GAL1* expression and chemoresistance (Chang et al. 2018; Cui et al. 2018), metastasis formation (Gessner et al. 1993) and poor prognosis (Lise et al. 2000), TCGA data did not support an association between *ST6GAL1* mRNA expression and clinical stage, response to therapy and overall survival. Mutation or overexpression of *RAS* has been associated with *ST6GAL1* overexpression in different cultured cell types (Dalziel et al. 2004; Seales et al. 2003), but the level of *ST6GAL1* mRNA was not higher in *KRAS* mutated specimens. The high *ST6GAL1*-expressing tumors displayed mainly high levels of the transcripts of *CEACAM5* and *CEACAM6*, two cell adhesion molecules of the CEA family promoting invasion and metastasis *in vitro* (Blumenthal et al. 2005). On the other hand, low *ST6GAL1*-expressing cases displayed up-regulation of a variety of genes involved in key features of malignancy, including carbohydrate metabolism, tissue repair, protein synthesis, cytoskeleton organization, and

mucin secretion. These associations do not necessarily imply a causal relationship between the expression of *ST6GAL1* and that of other genes. To study the causal effect of *ST6GAL1* overexpression on transcriptome modulation in different cellular contexts, we transduced with the *ST6GAL1* cDNA two cell lines, displaying the two main types of genomic instability in CRC. ST6GAL1-expressing cells exhibited higher α2,6-sialylation of different molecular weight proteins, although the effect on protein α2,6-sialylation appeared to be more pronounced in SW48 cells than in SW948 cells. On the other hand, the impact of ST6GAL1 on the transcriptome was stronger in SW948 than in SW48 cells. In fact, while in the MSS SW948 cells the expression of ST6GAL1 modulated numerous genes, a large number of which were putatively involved in cancer, in the MSI SW48 cell line only a few cancer-related genes displayed modulation. Notably, no genes revealed parallel and consistent ST6GAL1-induced modulation in the two cell lines, indicating that the influence of a glycosyltransferase on the transcriptional activity depends on the cellular context. A partial explanation for the differential transcriptomic response induced by ST6GAL1 in the two cell lines can be provided by the differential ability of ST6GAL1 to modulate the response to growth factors, as we documented for HGF.

Some phenotypic aspects of transduced cells could be explained by the transcriptional changes induced by ST6GAL1. For example, TGFB and WNT signaling predicted to be affected by ST6GAL1 expression in SW948 cells, could be responsible for the different ability to grow in soft agar and to heal a wound displayed by SW948 ST cells. Interestingly, while we show here that sialylation can affect TGFB expression, several lines of evidence show that TGFB signaling strongly impacts cell glycosylation and in particular sialylation (Lee et al. 2013; Lee et al. 2015; Lee et al. 2016). The increased ability to heal a scratch wound displayed by SW948 ST may depend both on their reduced ability to multilayer growth and from increased migration. This latter ability can be related to *MYEOV* overexpression, as documented for another CRC cell line (Lawlor et al. 2010). Regarding SW48 cells, the general ST6GAL1-induced down-regulation of *MAP1B* and of cytokeratin genes, which are involved in migration and cytoskeletal organization, could explain the reduced ability of SW48 ST to invade Matrigel. Altogether, these data indicate that the impact of ST6GAL1 expression on gene expression is potentially able to modulate multiple cancer-related cellular functions.

Other sialyltransferases have recently been shown to modulate transcriptional activity of cancer cell lines (Severino et al. 2017; Severino et al. 2018). This points to the mechanisms through which a glycosyltransferase, that encodes an enzyme acting mainly on membrane and secreted glycoproteins, can affect gene transcription. The most plausible mechanism implies that differentially glycosylated membrane receptors modulate signal transduction cascades (Gomes Ferreira et al. 2018), which

ultimately affect the transcription factor activity and the epigenetic pattern (Dall'Olio et al. 2017). Thus, the evident phenotypic changes in ST6GAL1-expressing cells can be explained by two partially overlapping mechanisms: a direct modulation of cell membrane receptors (for example the strength of integrin binding) and an indirect mechanism, through modulation of gene expression, which impacts multiple functions. Current results show that, whatever the changes induced by ST6GAL1 expression in colon cancer cells, they are strongly cell-type specific, ruling out the possibility that the association of ST6GAL1 and malignancy is a general paradigm. This should be kept into consideration for every future use of ST6GAL1 as a molecular target.

Materials and Methods

Analysis of transcriptomic databases

Gene expression data and clinical information for 623 colorectal adenocarcinoma samples and 51 normal tissues were downloaded from TCGA database using Firebrowse website (http://firebrowse.org). RSEM normalized data for colon adenocarcinoma ~~(COADREAD) cohort~~ were matched with clinical data from Clinical Pick Tier1 archive. Mutations for these patients were retrieved using cBioPortal web site http://www.cbioportal.org. ST6GAL1 mRNA expression was compared with stage, MSI status, response to treatment, histological type, survival, and *KRAS*, *BRAF*, *APC* and *TP53* mutations. Since the samples did not present a normal distribution of ST6GAL1 expression, non-parametric statistical tests were used. Mann Whitney test was used to analyze the difference of ST6GAL1 expression in normal and tumor tissues, mucinous vs. non-mucinous histological type, and in *KRAS, BRAF, APC* or *TP53* mutated vs. wild type patients. Kruskal-Wallis test was used to evaluate ST6GAL1 mRNA expression across cancer stages and MSS/MSI groups. The survival curve was estimated by the Kaplan-Meier method and the Mantel-Cox log-rank test was performed to test differences between the survival curves. Identification of highly expressed genes in the high and low ST6GAL1 expressers was performed through two-way ANOVA and Sidak's test for multiple comparisons. Gene expression in a different cohort of CRC cases (Ferracin et al. 2008) was analyzed by Agilent Technologies, and ST6GAL1 expression level was correlated with MSS or MSI-h status.

Cell lines

SW948 (ATCCR Number: CCL-237™) and SW48 (ATCCR Number: CCL-231™) cell line were cultured in Leibovitz's L-15 Medium in absence of CO_2 (L-15 is phosphate buffered) in a humidified incubator. Details on the lentiviral transduction with human *ST6GAL1* cDNA (ST cells) or with an empty vector (NC cells) were reported previously (Malagolini et al. 2009). Cell lines were genotyped

by using highly-polymorphic short tandem repeat loci, which were amplified using the PowerPlex® 16 HS System (Promega). Fragment analysis was done on an ABI3730xl (Life Technologies) and the resulting data were analyzed with GeneMarker HID software (Softgenetics) by Microsynth (Switzerland). Reports are available on request.

Transcriptomic analysis

Transcriptomic analysis of RNA from SW948 and SW48 NC and ST RNA was performed in duplicate using Agilent whole human genome oligo microarray (G4851A) as described (Ferracin et al. 2013). Statistical analysis was performed using two-way ANOVA for repeated measures and the false discovery rate was controlled with two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli with Q=0.05. Pathway analysis of differentially expressed genes was determined using the web-based software MetaCore (GeneGo, Thomson Reuters). Gene function was studied through an extensive literature search.

Real-Time PCR

Total RNA was extracted according to Chomczynski e Sacchi (Chomczynski et al. 1987) and reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) following manufacturer's instructions. The PCR reaction was carried out in triplicate with 20 ng cDNA, with TaqMan Fast Universal PCR Master Mix (Applied Biosystems), TaqMan probes for *ST6GAL1* (assay identification Hs00949382_m1) or *ACTB* (Hs99999903_m1) in a final volume of 10 µL.

HGF treatment

SW948 and SW48 NC and ST cells were plated at a concentration of 2 x 10^6 *per* well in a 6 multiwell plate, let adhere for 24-48 hours, serum-starved for 24 hours and treated with HGF at a final concentration of 25 ng/mL, for 15, 30 or 120 minutes. Cells were lysed for 15 minutes in ice with Ripa Buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1%NP-40, 0.5% DOC, 0.1% SDS supplemented with protease inhibitor and phosphatase inhibitor cocktail).

Western and SNA blot analysis

Equal amounts of proteins were electrophoresed according to Laemmli protocol in 8% acrylamide gels (or 6% for SNA-blot analysis) in denaturing and reducing conditions and blotted to activated PVDF membranes. Membranes were blocked for 1 hour at room temperature with 1.5% BSA in Tris-buffered saline and 0,05% Tween20 (TBS-T), then incubated overnight at 4°C with primary

antibody. After washing with TBS-T, membranes were incubated for 1 hour at room temperature with a secondary antibody conjugated with horseradish peroxidase (HRP). Detection was performed using Super Signal West Pico as a chemiluminescent substrate. Densitometric analysis was performed using Kodak 1D software and statistical analysis with two-way ANOVA followed by Holm-Sidak's multiple comparisons test. The following primary antibodies and lectins were used: mouse monoclonal anti-phosphotyrosine (Cytoskeleton) and anti-vinculin (Sigma), rabbit polyclonal anti-β-Actin (Sigma) and anti-FAK [pY397] (Life Technologies), goat polyclonal anti-ST6GAL1 (R&D systems). Digoxigenin-conjugated SNA was as described (Malagolini et al. 2009). As secondary antibodies, polyclonal rabbit anti-mouse (Sigma), goat anti-rabbit (Sigma), donkey anti-goat (R&D system) and anti-digoxigenin (Abcam) conjugated to HRP were used.

Soft agar growth assay

One mL/well of a 0.5% agar solution in complete L-15 was dispensed in each well of a 6 well plate and allowed to solidify. On the top of this layer of agar, one mL of a 0.3% agar solution in complete L-15 medium containing 1 $x10^4$ cells per well was dispensed in triplicate. The plates were incubated for two weeks at 37°C in a humidified incubator without CO_2. To evaluate the number of colonies formed, the plates were fixed and colored for one hour with a solution containing formaldehyde 4%and crystal violet 0.005% in PBS. Pictures were taken at 4 X magnification and colonies were counted. Statistical analysis was performed using non-parametric Kolmogorov-Smirnov test.

Wound healing assay

The wound healing assay was performed using Culture-Insert 2 Well (Ibidi). Aliquots of $7x10^4$ (for SW48) or $5x10^4$ (for SW948) cells were seeded in each well. When the cells reached confluency, the insert was removed and healing of the wound was measured by taking pictures at 4 X magnification. The area free of cells was measured using the MRI Wound Healing Tool in ImageJ (http://dev.mri.cnrs.fr/projects/imagej-macros/wiki/Wound_Healing_Tool). Statistical analysis was performed using two-way ANOVA and Tukey's multiple comparisons test.

Transwell invasion assay

The transwell invasion assay was performed using Matrigel-coated polycarbonate filters (8 µm pore size, Corning BioCoat Matrigel Invasion Chamber). Aliquots of $2x10^5$ (for SW948) or $4x10^5$ (for SW48) of 24 hours-serum starved cells were seeded in the upper chamber of the well in serum-free medium. Complete L-15 with 10% FBS was placed as chemoattractant in the lower chamber of the well. The plates were then incubated for 24 hours at 37°C in a humidified incubator without CO_2. Membranes were fixed in methanol, stained with toluidine blue, mounted on slides and cells were

counted with a microscope at a 10 X magnification. Statistical analysis was performed by using two-way ANOVA and Sidak's multiple comparisons test.

Stemness analysis

ALDEFLUOR (Stem Cell technologies) was activated following manufacturer's instructions and added to 2 x 10^5 aliquots of cells. Half of the cell suspension was treated with DEAB, a specific ALDH inhibitor used for negative control. After 30 minutes at 37°C, cells were washed and suspended in ALDEFLUOR buffer. The fluorescent signal was acquired with a FACSCalibur flow cytometer and Cell Quest Pro software. On a dot plot with FL1 (green fluorescence) on the X axis and SSC on the Y axis, we set the fluorescence of the DEAB sample (negative control) and defined the area for ALDH positive cells. Cells included in this area were considered ALDEFLUOR positive. An aliquot of 3 x 10^5 cells was incubated with mouse monoclonal CD133/1 (AC133)-phycoerythrin antibody (Miltenyi Biotech) for 10 minutes at 4°C in the dark in the appropriate buffer (PBS pH 7.2, 0.5% BSA and 2 mM EDTA). After a wash, cells were analyzed by flow cytometry.

Acknowledgments

We would like to thank the LTTA Microarray Facility of the University of Ferrara (Italy) for performing the microarray experiments.

Conflict of Interest statement

None declared.

Funding

This work was supported by the European Commission Horizon 2020 program under grant agreement number 676421 (GlyCoCan), by funds from the University of Bologna and by the "Pallotti" Legacy for Cancer Research to FDO.

Abbreviations:

ALDH: aldheyde dehydrogenase 1; ANOVA: Analysis of Variance; BRAF: Rapidly Accelerated Fibrosarcoma B; BSA: bovine serum albumin; CIN: chromosomal instability; CRC: Colorectal cancer; DEAB: N,N-diethylaminobenzaldehyde(ALDH inhibitor); EMT: epithelial to mesenchymal transition; FAK: focal adhesion kinase; FBS: fetal bovine serum; FITC: Fluorescein isothiocyanate; HGF: hepatocyte growth factor; MSI microsatellite instability; MSI-h: high microsatellite

instability; MSI-l: low microsatellite instability; MSS: microsatellite stability; PBS: phosphate-buffered saline; PCR: polymerase chain reaction; SNA: *Sambucus nigra* agglutinin; TCGA: The Cancer Genome Atlas; TGFB:transforming growth factor-β.

References

Antony P, Rose M, Heidenreich A, Knuchel R, Gaisa NT, Dahl E. 2014. Epigenetic inactivation of ST6GAL1 in human bladder cancer. *BMC Cancer*. 14:901

Blumenthal RD, Hansen HJ, Goldenberg DM. 2005. Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen). *Cancer Res*. 65:8809-8817

Britain CM, Dorsett KA, Bellis SL. 2017. The Glycosyltransferase ST6Gal-I Protects Tumor Cells against Serum Growth Factor Withdrawal by Enhancing Survival Signaling and Proliferative Potential. *J Biol Chem*. 292:4663-4673

Britain CM, Holdbrooks AT, Anderson JC, Willey CD, Bellis SL. 2018. Sialylation of EGFR by the ST6Gal-I sialyltransferase promotes EGFR activation and resistance to gefitinib-mediated cell death. *J Ovarian Res*. 11:12

Chakraborty A, Dorsett KA, Trummell HQ, Yang ES, Oliver PG, Bonner JA, Buchsbaum DJ, Bellis SL. 2018. ST6Gal-I sialyltransferase promotes chemoresistance in pancreatic ductal adenocarcinoma by abrogating gemcitabine-mediated DNA damage. *J Biol Chem*. 293:984-994

Chang TC, Chin YT, Nana AW, Wang SH, Liao YM, Chen YR, Shih YJ, Changou CA, Yang YS, Wang K, Whang-Peng J, Wang LS, Stain SC, Shih A, Lin HY, Wu CH, Davis PJ. 2018. Enhancement by Nano-Diamino-Tetrac of Antiproliferative Action of Gefitinib on Colorectal Cancer Cells: Mediation by EGFR Sialylation and PI3K Activation. *Horm Cancer*. 9:420-432

Chiricolo M, Malagolini N, Bonfiglioli S, Dall'Olio F. 2006. Phenotypic changes induced by expression of β-galactoside α2,6 sialyltransferase I in the human colon cancer cell line SW948. *Glycobiology*. 16:146-154

Chomczynski P Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal Biochem*. 162:156-159

Cui H, Yang S, Jiang Y, Li C, Zhao Y, Shi Y, Hao Y, Qian F, Tang B, Yu P. 2018. The glycosyltransferase ST6Gal-I is enriched in cancer stem-like cells in colorectal carcinoma and contributes to their chemo-resistance. *Clin Transl Oncol*. 20:1175-1184

Dall'Olio F. 2000. The sialyl-α2,6-lactosaminyl-structure: biosynthesis and functional role. *Glycoconj J*. 17:669-676

Dall'Olio F Chiricolo M. 2001. Sialyltransferases in cancer. *Glycoconj J*. 18:841-850

Dall'Olio F, Chiricolo M, Lollini P, Lau JT. 1995. Human colon cancer cell lines permanently expressing α2,6- sialylated sugar chains by transfection with rat β-galactoside α 2,6 sialyltransferase cDNA. *Biochem Biophys Res Commun*. 211:554-561

Dall'Olio F, Malagolini N, Di Stefano G, Minni F, Marrano D, Serafini-Cessi F. 1989. Increased CMP-NeuAc:Galβ1,4GlcNAc-R α 2,6 sialyltransferase activity in human colorectal cancer tissues. *Int J Cancer*. 44:434-439

Dall'Olio F, Malagolini N, Trinchera M, Chiricolo M. 2012. Mechanisms of cancer-associated glycosylation changes. *Front Biosci*. 17:670-699

Dall'Olio F, Malagolini N, Trinchera M, Chiricolo M. 2014. Sialosignaling: Sialyltransferases as engines of self-fueling loops in cancer progression. *Biochim Biophys Acta*. 1840:2752-2764

Dall'Olio F Trinchera M. 2017. Epigenetic Bases of Aberrant Glycosylation in Cancer. *Int J Mol Sci*. 18:

Dalziel M, Dall'Olio F, Mungul A, Piller V, Piller F. 2004. Ras oncogene induces β-galactoside α2,6-sialyltransferase (ST6Gal I) via a RalGEF-mediated signal to its housekeeping promoter. *Eur J Biochem*. 271:3623-3634

Dawson G, Moskal JR, Dawson SA. 2004. Transfection of 2,6 and 2,3-sialyltransferase genes and GlcNAc-transferase genes into human glioma cell line U-373 MG affects glycoconjugate expression and enhances cell death. *J Neurochem*. 89:1436-1444

Ferracin M, Bassi C, Pedriali M, Pagotto S, D'Abundo L, Zagatti B, Corra F, Musa G, Callegari E, Lupini L, Volpato S, Querzoli P, Negrini M. 2013. miR-125b targets erythropoietin and its receptor and their expression correlates with metastatic potential and ERBB2/HER2 expression. *Mol Cancer*. 12:130

Ferracin M, Gafa R, Miotto E, Veronese A, Pultrone C, Sabbioni S, Lanza G, Negrini M. 2008. The methylator phenotype in microsatellite stable colorectal cancers is characterized by a distinct gene expression profile. *J Pathol*. 214:594-602

Fukushima K, Hara-Kuge S, Seko A, Ikehara Y, Yamashita K. 1998. Elevation of α2,6 sialyltransferase and α1,2 fucosyltransferase activities in human choriocarcinoma. *Cancer Res*. 58:4301-4306

Gebert J, Kloor M, Lee J, Lohr M, Andre S, Wagner R, Kopitz J, Gabius HJ. 2012. Colonic carcinogenesis along different genetic routes: glycophenotyping of tumor cases separated by microsatellite instability/stability. *Histochem Cell Biol*. 138:339-350

Gessner P, Riedl S, Quentmaier A, Kemmner W. 1993. Enhanced activity of CMP-NeuAc:Gal β 1-4GlcNAc:α 2,6-sialyltransferase in metastasizing human colorectal tumor tissue and serum of tumor patients. *Cancer Lett*. 75:143-149

Gomes Ferreira I, Pucci M, Venturi G, Malagolini N, Chiricolo M, Dall'Olio F. 2018. Glycosylation as a Main Regulator of Growth and Death Factor Receptors Signaling. *Int J Mol Sci*. 19:E580

Han Y, Liu Y, Fu X, Zhang Q, Huang H, Zhang C, Li W, Zhang J. 2018. miR-9 inhibits the metastatic ability of hepatocellular carcinoma via targeting beta galactoside alpha-2,6-sialyltransferase 1. *J Physiol Biochem*.

Jones RB, Dorsett KA, Hjelmeland AB, Bellis SL. 2018. The ST6Gal-I sialyltransferase protects tumor cells against hypoxia by enhancing HIF-1alpha signaling. *J Biol Chem*. 253:5659-5667

Jun L, Yuanshu W, Yanying X, Zhongfa X, Jian Y, Fengling W, Xianjun Q, Kokudo N, Wei T, Weixia Z, Shuxiang C. 2012. Altered mRNA expressions of sialyltransferases in human gastric cancer tissues. *Med Oncol*. 29:84-90

Jung YR, Park JJ, Jin YB, Cao YJ, Park MJ, Kim EJ, Lee M. 2016. Silencing of ST6Gal I enhances colorectal cancer metastasis by down-regulating KAI1 via exosome-mediated exportation and thereby rescues integrin signaling. *Carcinogenesis*. 37:1089-1097

Lai JF, Kao SC, Jiang ST, Tang MJ, Chan PC, Chen HC. 2000. Involvement of focal adhesion kinase in hepatocyte growth factor-induced scatter of Madin-Darby canine kidney cells. *J Biol Chem*. 275:7474-7480

Lawlor G, Doran PP, MacMathuna P, Murray DW. 2010. MYEOV (myeloma overexpressed gene) drives colon cancer cell migration and is regulated by PGE2. *J Exp Clin Cancer Res*. 29:81

Lee J, Ballikaya S, Schonig K, Ball CR, Glimm H, Kopitz J, Gebert J. 2013. Transforming growth factor beta receptor 2 (TGFBR2) changes sialylation in the microsatellite unstable (MSI) Colorectal cancer cell line HCT116. *PLoS One*. 8:e57074

Lee J, Katzenmaier EM, Kopitz J, Gebert J. 2016. Reconstitution of TGFBR2 in HCT116 colorectal cancer cells causes increased LFNG expression and enhanced N-acetyl-d-glucosamine incorporation into Notch1. *Cell Signal*. 28:1105-1113

Lee J, Warnken U, Schnolzer M, Gebert J, Kopitz J. 2015. A new method for detection of tumor driver-dependent changes of protein sialylation in a colon cancer cell line reveals nectin-3 as TGFBR2 target. *Protein Sci*. 24:1686-1694

Lee M, Lee HJ, Bae S, Lee YS. 2008. Protein sialylation by sialyltransferase involves radiation resistance. *Mol Cancer Res*. 6:1316-1325

Lee M, Lee HJ, Seo WD, Park KH, Lee YS. 2010a. Sialylation of integrin β1 is involved in radiation-induced adhesion and migration in human colon cancer cells. *Int J Radiat Oncol Biol Phys*. 76:1528-1536

Lee M, Park JJ, Lee YS. 2010b. Adhesion of ST6Gal I-mediated human colon cancer cells to fibronectin contributes to cell survival by integrin β1-mediated paxillin and AKT activation. *Oncol Rep*. 23:757-761

Lise M, Belluco C, Perera SP, Patel R, Thomas P, Ganguly A. 2000. Clinical correlations of α2,6-sialyltransferase expression in colorectal cancer patients. *Hybridoma*. 19:281-286

Malagolini N, Chiricolo M, Marini M, Dall'Olio F. 2009. Exposure of α2,6-sialylated lactosaminic chains marks apoptotic and necrotic death in different cell types. *Glycobiology*. 19:172-181

Meng Q, Ren C, Wang L, Zhao Y, Wang S. 2015. Knockdown of ST6Gal-I inhibits the growth and invasion of osteosarcoma MG-63 cells. *Biomed Pharmacother*. 72:172-178

Mondal S, Chandra S, Mandal C. 2010. Elevated mRNA level of hST6Gal I and hST3Gal V positively correlates with the high risk of pediatric acute leukemia. *Leuk Res*. 34:463-470

Park JJ, Yi JY, Jin YB, Lee YJ, Lee JS, Lee YS, Ko YG, Lee M. 2012. Sialylation of epidermal growth factor receptor regulates receptor activity and chemosensitivity to gefitinib in colon cancer cells. *Biochem Pharmacol*. 83:849-857

Petretti T, Kemmner W, Schulze B, Schlag PM. 2000. Altered mRNA expression of glycosyltransferases in human colorectal carcinomas and liver metastases. *Gut*. 46:359-366

Pinho SS Reis CA. 2015. Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer*. 15:540-555

Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, De Maria R. 2007. Identification and expansion of human colon-cancer-initiating cells. *Nature*. 445:111-115

Schultz MJ, Holdbrooks AT, Chakraborty A, Grizzle WE, Landen CN, Buchsbaum DJ, Conner MG, Arend RC, Yoon KJ, Klug CA, Bullard DC, Kesterson RA, Oliver PG, O'Connor AK, Yoder BK, Bellis SL. 2016. The Tumor-Associated Glycosyltransferase ST6Gal-I Regulates Stem Cell Transcription Factors and Confers a Cancer Stem Cell Phenotype. *Cancer Res*. 76:3978-3988

Seales EC, Jurado GA, Brunson BA, Wakefield JK, Frost AR, Bellis SL. 2005. Hypersialylation of β1 integrins, observed in colon adenocarcinoma, may contribute to cancer progression by up-regulating cell motility. *Cancer Res*. 65:4645-4652

Seales EC, Jurado GA, Singhal A, Bellis SL. 2003. Ras oncogene directs expression of a differentially sialylated, functionally altered β1 integrin. *Oncogene*. 22:7137-7145

Severino PF, Silva M, Carrascal M, Malagolini N, Chiricolo M, Venturi G, Astolfi A, Catera M, Videira PA, Dall'Olio F. 2017. Expression of sialyl-Tn sugar antigen in bladder cancer cells

affects response to Bacillus Calmette Guerin (BCG) and to oxidative damage. *Oncotarget*. 8:54506-54517

Severino PF, Silva M, Carrascal M, Malagolini N, Chiricolo M, Venturi G, Barbaro FR, Astolfi A, Catera M, Videira PA, Dall'Olio F. 2018. Oxidative damage and response to Bacillus Calmette-Guerin in bladder cancer cells expressing sialyltransferase ST3GAL1. *BMC Cancer*. 18:198

Shah MH, Telang SD, Shah PM, Patel PS. 2008. Tissue and serum α2-3- and α2-6-linkage specific sialylation changes in oral carcinogenesis. *Glycoconj J*. 25:279-290

Shibuya N, Goldstein IJ, Broekaert WF, Nsimba-Lubaki M, Peeters B, Peumans WJ. 1987. The elderberry (*Sambucus nigra* L.) bark lectin recognizes the Neu5Ac(α 2-6)Gal/GalNAc sequence. *J Biol Chem*. 262:1596-1601

Skacel PO, Edwards AJ, Harrison CT, Watkins WM. 1991. Enzymic control of the expression of the X determinant (CD15) in human myeloid cells during maturation: the regulatory role of 6-sialytransferase. *Blood*. 78:1452-1460

Swindall AF Bellis SL. 2011. Sialylation of the Fas Death Receptor by ST6Gal-I Provides Protection against Fas-mediated Apoptosis in Colon Carcinoma Cells. *J Biol Chem*. 286:22982-22990

Swindall AF, Londono-Joshi AI, Schultz MJ, Fineberg N, Buchsbaum DJ, Bellis SL. 2013. ST6Gal-I Protein Expression Is Upregulated in Human Epithelial Tumors and Correlates with Stem Cell Markers in Normal Tissues and Colon Cancer Cell Lines. *Cancer Res*. 73:2368-2378

Vierbuchen MJ, Fruechtnicht W, Brackrock S, Krause KT, Zienkiewicz TJ. 1995. Quantitative lectin-histochemical and immunohistochemical studies on the occurrence of α(2,3)- and α(2,6)-linked sialic acid residues in colorectal carcinomas. Relation to clinicopathologic features. *Cancer*. 76:727-735

Volonte A, Di TT, Spinelli M, Todaro M, Sanvito F, Albarello L, Bissolati M, Ghirardelli L, Orsenigo E, Ferrone S, Doglioni C, Stassi G, Dellabona P, Staudacher C, Parmiani G, Maccalli C. 2014. Cancer-initiating cells from colorectal cancer patients escape from T cell-mediated immunosurveillance in vitro through membrane-bound IL-4. *J Immunol*. 192:523-532

Wang YC, Stein JW, Lynch CL, Tran HT, Lee CY, Coleman R, Hatch A, Antontsev VG, Chy HS, O'Brien CM, Murthy SK, Laslett AL, Peterson SE, Loring JF. 2015. Glycosyltransferase ST6GAL1 contributes to the regulation of pluripotency in human pluripotent stem cells. *Sci Rep*. 5:13317

Wei A, Fan B, Zhao Y, Zhang H, Wang L, Yu X, Yuan Q, Yang D, Wang S. 2016. ST6Gal-I overexpression facilitates prostate cancer progression via the PI3K/Akt/GSK-3β/β-catenin signaling pathway. *Oncotarget*. 7:65374-65388

Weinstein J, Lee EU, McEntee K, Lai PH, Paulson JC. 1987. Primary structure of β-galactoside α 2,6-sialyltransferase. Conversion of membrane-bound enzyme to soluble forms by cleavage of the NH_2-terminal signal anchor. *J Biol Chem*. 262:17735-17743

Yamamoto H, Kaneko Y, Rebbaa A, Bremer EG, Moskal JR. 1997. α2,6-Sialyltransferase gene transfection into a human glioma cell line (U373 MG) results in decreased invasivity. *J Neurochem*. 68:2566-2576

Yamamoto H, Oviedo A, Sweeley C, Saito T, Moskal JR. 2001. α2,6-Sialylation of cell-surface N-glycans inhibits glioma formation in vivo. *Cancer Res*. 61:6822-6829

Yamashita K, Fukushima K, Sakiyama T, Murata F, Kuroki M, Matsuoka Y. 1995. Expression of Sia α 2-->6Gal β 1-->4GlcNAc residues on sugar chains of glycoproteins including

carcinoembryonic antigens in human colon adenocarcinoma: applications of *Trichosanthes japonica* agglutinin I for early diagnosis. *Cancer Res*. 55:1675-1679

Yen HY, Liu YC, Chen NY, Tsai CF, Wang YT, Chen YJ, Hsu TL, Yang PC, Wong CH. 2015. Effect of sialylation on EGFR phosphorylation and resistance to tyrosine kinase inhibition. *Proc Natl Acad Sci U S A*. 112:6955-6960

Zhang S, Lu J, Xu Z, Zou X, Sun X, Xu Y, Shan A, Lu J, Yan X, Cui Y, Yan W, Du Y, Gu J, Zheng M, Feng B, Zhang Y. 2017. Differential expression of ST6GAL1 in the tumor progression of colorectal cancer. *Biochem Biophys Res Commun*. 486:1090-1096

Zhuo Y, Chammas R, Bellis SL. 2008. Sialylation of β1 integrins blocks cell adhesion to galectin-3 and protects cells against galectin-3-induced apoptosis. *J Biol Chem*. 283:22177-22185

Legend to Figures

Fig. I. Data from the TCGA database. A: *ST6GAL1* mRNA expression in cancer tissues of CRC patients. B-J: box plot graphs showing median and Q1 and Q3 quartiles of *ST6GAL1* mRNA expression in CRC and normal tissues (B); in CRC tissue of stage I-IV patients (C); in CRC tissue of patients with microsatellite stability (MSS), low grade microsatellite instability (MSI-l), high grade microsatellite instability (MSI-h) (D); in CRC tissue of patients with wild type or mutated *APC* (E), *TP53* (F), *KRAS* (G), *BRAF* (H); with mucinous or non-mucinous phenotype (I); showing response or no-response to therapy (J). K: Survival curves of patients falling in the upper or lower 15% percentile of *ST6GAL1* mRNA expression (93 cases). Mann Whitney test was used in B, E, F, G, H, I, J. Kruskal-Wallis test was used in C and D. The Mantel-Cox test was used in K. **** $p \leq 0.0001$

Fig. II. Heatmaps of gene expression of *ST6GAL1*-transduced SW948 and SW48 cells. The genes that are differentially expressed between ST6GAL1-transduced cells (ST) and respective negative control NC are reported. 1 and 2 refer to the two independent replicates. Genes (columns) and samples (rows) were grouped by hierarchical clustering (Manhattan correlation). High- and low- expression was normalized to the average expression across all samples. To analyze the differences between *ST6GAL1*-transduced and negative control cells the moderated *t*-test was used. Corrected *p*-value cut-off: 0.15; multiple test correction used:Benjamini-Hochberg.

Fig. III. HGF-induced tyrosine phosphorylation in ST6GAL1-transduced SW948 and SW48 cells. Cells were treated with HGF for the indicated times, harvested as described in Materials and Methods and analyzed by Western blot for phosphorylation of Tyr397 of FAK (upper panels) or for phosphotyrosines in SW948 (A) or SW48 (B). The chemiluminescent signals were normalized with that of β-actin. Histograms on the right represent the mean±SD of three independent experiments. *$p<0.05$.

Fig. IV. Phenotypic characterization of *ST6GAL1*-transduced SW948 and SW48 cells. A: Soft agar growth. Mock- or *ST6GAL1*-transduced cells were seeded in soft agar and the number of the colonies was determined as described in Materials and Methods. Representative fields are shown on the left (original magnification 5 X). The total number of colonies is reported on the right. Statistical analysis was performed using non-parametric Kolmogorov-Smirnov test. B: Scratch wound test. A wound was made in a confluent cell layer, as described in Material and Methods. The healing of the wound was monitored by taking pictures at given time intervals. Representative fields, taken at the time of the wound (Day 0) or seven days later (Day 7), are shown at the top (original magnification 10 X), while the percentage of free wound measured after seven days of healing is reported at the

bottom. Statistical analysis was performed using two-way ANOVA and Tukey's multiple comparisons test. C: Matrigel invasion test. Cells were seeded in the upper chamber of the transwell in serum-free medium. Complete medium was placed in the lower chamber of the well as chemoattractant. The membranes were fixed, stained, mounted on slides and the cells were counted using a microscope at a 10 X magnification. Statistical analysis was performed using two-way ANOVA and Sidak's multiple comparisons test. Photographs show representative experiments. Histograms report mean±SD of three independent experiments. ** $p<0.01$; **** $p<0.0001$.

Fig. V. Stem cell markers in *ST6GAL1*-transduced SW948 and SW48 cells. A: Cells were incubated with ALDEFLUOR either in the presence or in the absence of DEAB, a specific ALDH inhibitor used for negative control, as described in Materials and Methods. A gate excluding the vast majority of the cells labelled in the presence of DEAB was set (upper panels). Cells included in the gate in the absence of DEAB, were considered to be ALDH-positive. Numbers indicate the percentage of ALDH positive cells ±SD in three independent experiments. B: Cells were treated with anti-CD133 antibodies as detailed in Material and Methods and FACS analyzed. Red: NC cells, Blue: ST cells. The numbers indicate the percentages of cells included in the sections $CD133^{-}$, $CD133^{+}$ and $CD133^{+high}$.

Table I. Genes up-regulated in high- or low *ST6GAL1* expressers, according to TCGA database

Genes up-regulated in high *ST6GAL1* expressers			
Gene name	**Protein**	**Functions**	**Functional class**
CD24	**CD24** CD24 protein	Sialoglycoprotein cell adhesion molecule. Promotes adhesion and metastasis	Cell adhesion
CEACAM5	Carcinoembryonic Antigen Related Cell Adhesion Molecule 5	Member of the CEA family. Biomarker for gastrointestinal cancers. Promotes cell adhesion and metastasis.	
CEACAM6	Carcinoembryonic Antigen Related Cell Adhesion Molecule 5	Member CEA family. Promotes cell adhesion and metastasis.	
FTL	Ferritin light chain	Iron storage	Iron metabolism
PYGB	Glycogen Phosphorylase B	Catalyzes the rate-determining step in glycogen degradation	Carbohydrate metabolism
Genes up-regulated in low *ST6GAL1* expressers			
PKM2	Muscle pyruvate kinase	Contributes to the control of glycolysis and is important for tumor cell proliferation and survival	Carbohydrate metabolism
GAPDH	Glyceraldhyde-3-phosphate kinase	Multiple functions	
REG1A	Regenerating family member 1α	Regenerating proteins are acute phase reactants, lectins, antiapoptotic or growth factors.	Regeneration and repair
TFF3	Trefoil factor 3	Involved in the maintenance and repair of the intestinal mucosa. Promotes the mobility of epithelial cells in healing processes	
RPL8	Ribosomal protein L8	Component of the 60S ribosomal subunit	Protein synthesis Protein synthesis
RPLP0	Ribosomal protein lateral stalk subunit P0		
EEF1A1	Eukariotic translation elongation factor 1α1	Promotes the binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis	
EEF2	Eukariotic translation elongation factor 2	Member of the GTP-binding translation elongation factor family essential for protein synthesis	
SERPINA1	Serpin family A member 1	Serine protease inhibitor for elastase, plasmin, thrombin, trypsin, chymotrypsin, and plasminogen activator	Protease inhibitor
LYZ	Lysozime	Bacteriolytic and enhancer of the activity of immunoagents	Antimicrobial and immune function

HLA-C	MHC class 1, C	Presentation of foreign antigens to CD8+ lymphocytes	
ACTB	Actin beta	Cytoskeleton component	Cytoskeleton
ACTG1	Actin gamma 1	A cytoplasmic actin found in non-muscle cells	
KRT18	Keratin 18	Member of the intermediate filament family. Partner of KRT8	
KRT19	Keratin 19	Member of the intermediate filament family	
KRT8	Keratin 8	Partner of KRT18 in intermediate filaments	
ANXA2	Annexin A2	Involved in cell motility, organization of cytoskeleton and exocytosis	
S100A6	S100 calcium binding protein A6	Indirectly plays a role in the reorganization of the actin cytoskeleton and in cell motility	
MUC2	Mucin 2, Oligomeric Mucus/Gel-Forming	Coats the epithelia of the intestines and other mucous membranes. Downreguled in IBDs	Mucin-related
MUC5B	Mucin 5B, Oligomeric Mucus/Gel-Forming	Gel-forming mucin that contributes to the lubricating and viscoelastic properties of mucus	
AGR2	Anterior Gradient 2, Protein Disulphide Isomerase Family Member	Required for MUC2 post-transcriptional synthesis and secretion. Proto-oncogene	
CLCA1	loride Channel Accessory 1	Involved in chloride conductance and regulation of mucus production. Potential tumor suppressor. Induces MUC5AC	
FCGBP	Fc Fragment of IgG Binding Protein	May be involved in the maintenance of the mucosal structure as a gel-like component of the mucosa	
COL1A1	Collagen Type I Alpha 1 Chain	Pro-alpha1 chains of type I collagen	Extracellular matrix

Statistical analysis was performed using Two-way ANOVA and Sidak's test for multiple comparisons.

Table II. CRC cellular pathways enriched by ST6GAL1 overexpression in SW948 and SW48 cells

SW948			
Pathways and cellular processes	**Number of involved genes**	***p* value**	**Genes in the pathway**
Up-regulation of IL-8 in CRC	2/32	0.0055	WNT4, WNT
TGF-β-dependent induction of EMT via SMADs	2/32	0.0065	TGF-β, TGF-β2
Modulation of Ca^{++} and K^{+} channels by Adenosine A1 receptor in signal transduction	2/43	0.0097	SUR1, Kir6.2
Cytoskeleton remodeling. Regulation of actin cytoskeleton nucleation and polymerization by Rho GTPases	2/46	0.011	FMNL1, DRF
Leptin signaling via PI3K-dependent pathway	2/47	0.011	SUR1, Kir6.2
TGF-β-dependent induction of EMT via MAPK	2/47	0.011	TGF-β, TGF-β2
Canonical WNT signaling pathway in CRC	2/66	0.022	WNT4, WNT
Transcription of HIF-1 targets	2/95	0.043	TGF-β2, SDF-1
SW48			
Estradiol metabolism	3/42	0.0002	UGT2B11, UGT2B7, UGT2B28
Inhibition of Ephrin receptors in CRC	2/30	0.0031	Ephrin-A receptors, Ephrin-A receptor 3
Cytoskeleton remodelling	3/36	0.0044	Keratin 14, Keratin 16, Keratin 17
Androsterone and testosterone biosynthesis	2/40	0.005	UGT2B15, UGT2B7, UGT2B28

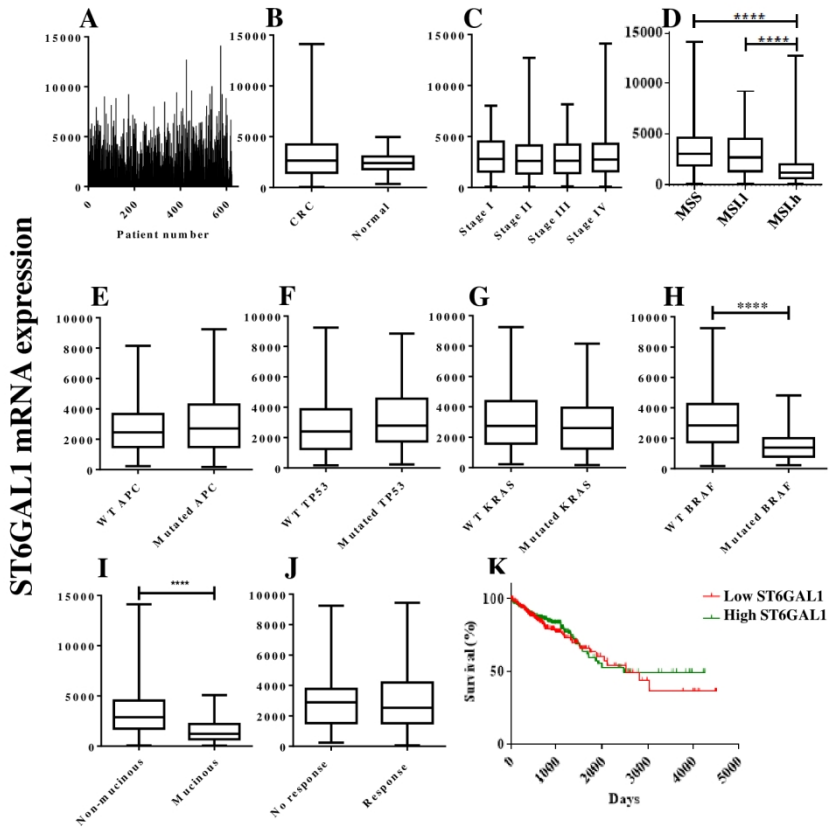

Fig. 1

Fig. I. Data from the TCGA database. A: ST6GAL1 mRNA expression in cancer tissues of CRC patients. B-J: box plot graphs showing median and Q1 and Q3 quartiles of ST6GAL1 mRNA expression in CRC and normal tissues (B); in CRC tissue of stage I-IV patients (C); in CRC tissue of patients with microsatellite stability (MSS), low grade microsatellite instability (MSI-l), high grade microsatellite instability (MSI-h) (D); in CRC tissue of patients with wild type or mutated APC (E), TP53 (F), KRAS (G), BRAF (H); with mucinous or non-mucinous phenotype (I); showing response or no-response to therapy (J). K: Survival curves of patients falling in the upper or lower 15% percentile of ST6GAL1 mRNA expression (93 cases). Mann Whitney test was used in B, E, F, G, H, I, J. Kruskal-Wallis test was used in C and D. The Mantel-Cox test was used in K. **** p≤0.0001

190x275mm (200 x 200 DPI)

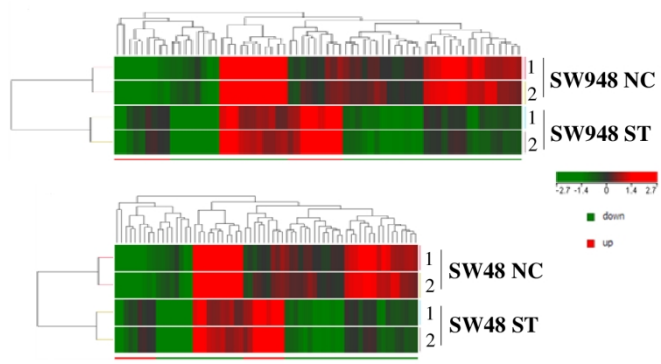

Fig. 2

Fig. II. Heatmaps of gene expression of ST6GAL1-transduced SW948 and SW48 cells. The genes that are differentially expressed between ST6GAL1-transduced cells (ST) and respective negative control NC are reported. 1 and 2 refer to the two independent replicates. Genes (columns) and samples (rows) were grouped by hierarchical clustering (Manhattan correlation). High- and low- expression was normalized to the average expression across all samples. To analyze the differences between ST6GAL1-transduced and negative control cells the moderated t-test was used. Corrected p-value cut-off: 0.15; multiple test correction used:Benjamini-Hochberg.

254x190mm (200 x 200 DPI)

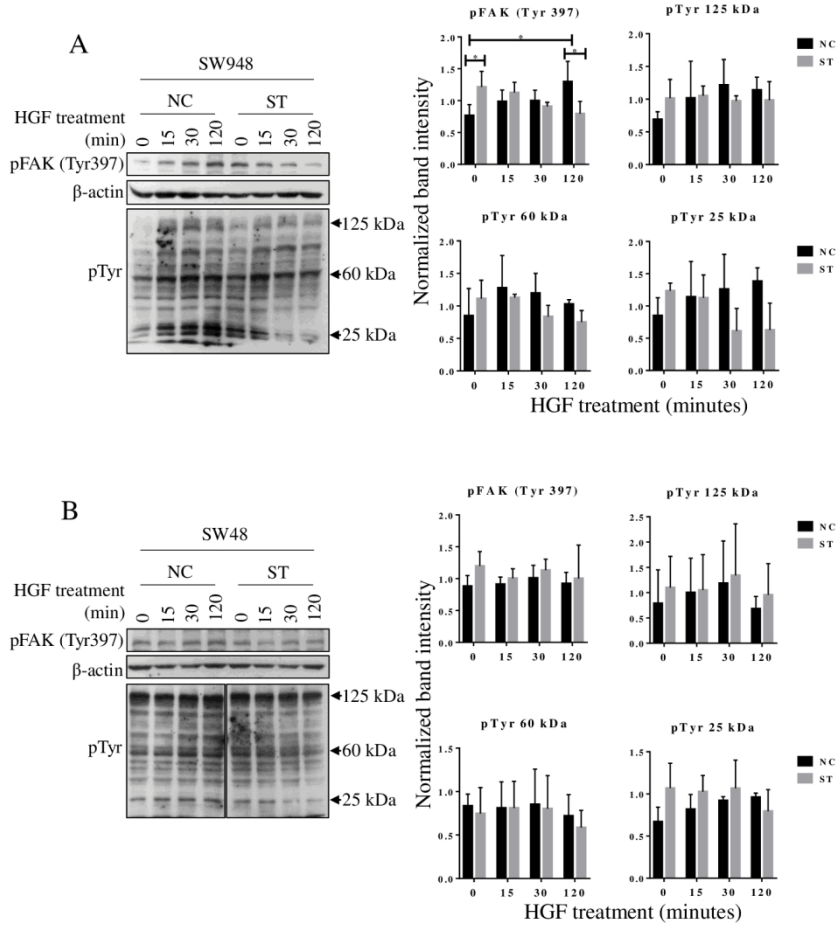

Fig. 3

Fig. III. HGF-induced tyrosine phosphorylation in ST6GAL1-transduced SW948 and SW48 cells. Cells were treated with HGF for the indicated times, harvested as described in Materials and Methods and analyzed by Western blot for phosphorylation of Tyr397 of FAK (upper panels) or for phosphotyrosines in SW948 (A) or SW48 (B). The chemiluminescent signals were normalized with that of β-actin. Histograms on the right represent the mean±SD of three independent experiments. *p<0.05.

190x275mm (200 x 200 DPI)

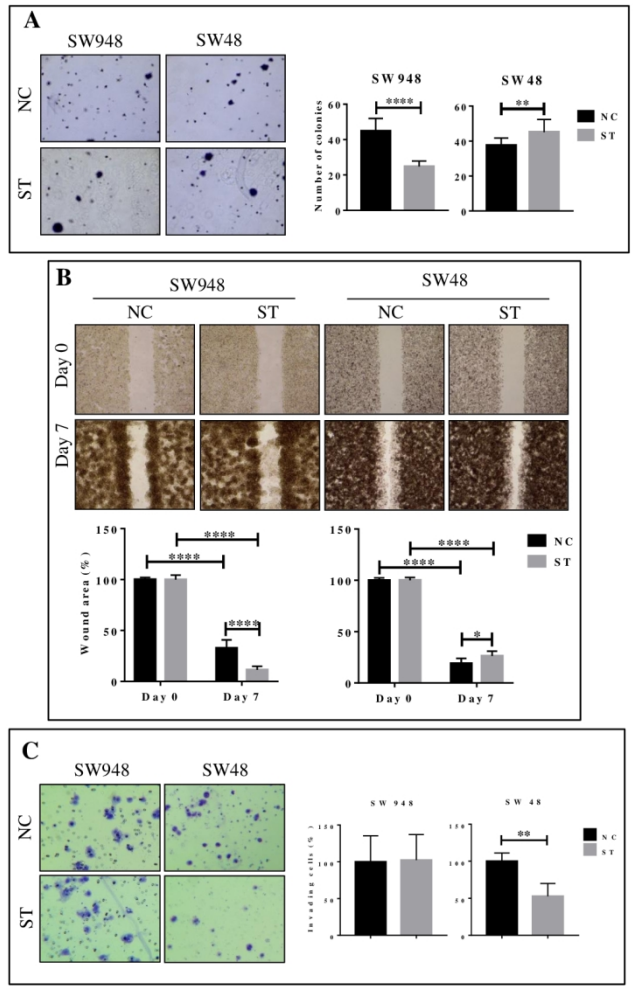

Fig. 4

Fig. IV. Phenotypic characterization of ST6GAL1-transduced SW948 and SW48 cells. A: Soft agar growth. Mock- or ST6GAL1-transduced cells were seeded in soft agar and the number of the colonies was determined as described in Materials and Methods. Representative fields are shown on the left (original magnification 5 X). The total number of colonies is reported on the right. Statistical analysis was performed using non-parametric Kolmogorov-Smirnov test. B: Scratch wound test. A wound was made in a confluent cell layer, as described in Material and Methods. The healing of the wound was monitored by taking pictures at given time intervals. Representative fields, taken at the time of the wound (Day 0) or seven days later (Day 7), are shown at the top (original magnification 10 X), while the percentage of free wound measured after seven days of healing is reported at the bottom. Statistical analysis was performed using two-way ANOVA and Tukey's multiple comparisons test. C: Matrigel invasion test. Cells were seeded in the upper chamber of the transwell in serum-free medium. Complete medium was placed in the lower chamber of the well as chemoattractant. The membranes were fixed, stained, mounted on slides and the cells were counted using a microscope at a 10 X magnification. Statistical analysis was performed using two-way ANOVA and Sidak's multiple comparisons test. Photographs show representative experiments. Histograms report

mean±SD of three independent experiments. ** p<0.01; **** p<0.0001.

190x275mm (200 x 200 DPI)

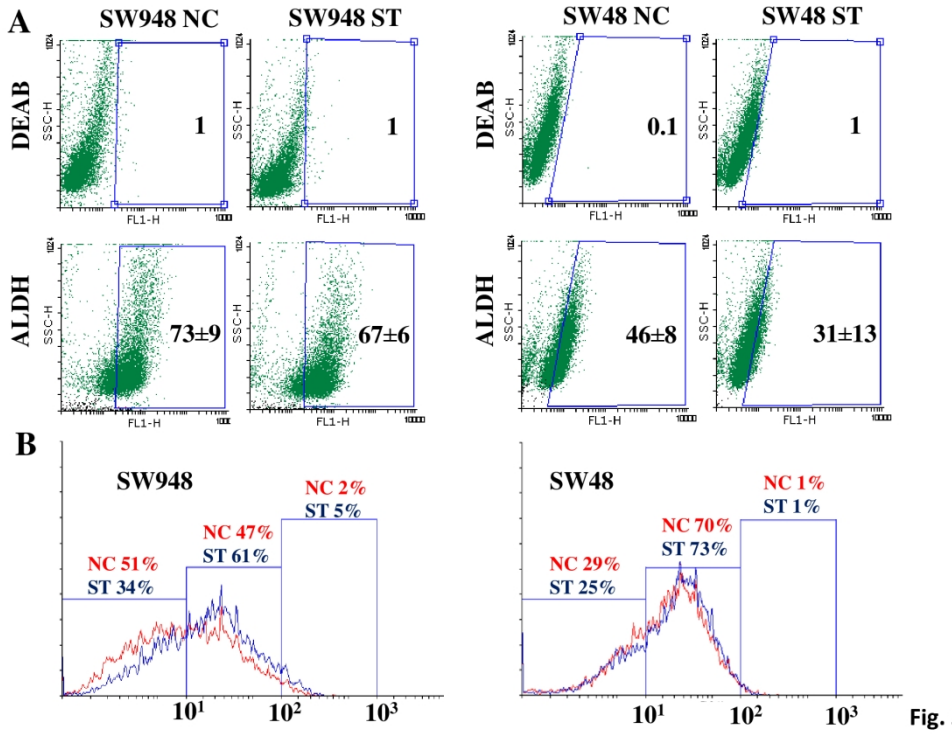

Fig. V. Stem cell markers in ST6GAL1-transduced SW948 and SW48 cells. A: Cells were incubated with ALDEFLUOR either in the presence or in the absence of DEAB, a specific ALDH inhibitor used for negative control, as described in Materials and Methods. A gate excluding the vast majority of the cells labelled in the presence of DEAB was set (upper panels). Cells included in the gate in the absence of DEAB, were considered to be ALDH-positive. Numbers indicate the percentage of ALDH positive cells ±SD in three independent experiments. B: Cells were treated with anti-CD133 antibodies as detailed in Material and Methods and FACS analyzed. Red: NC cells, Blue: ST cells. The numbers indicate the percentages of cells included in the sections CD133-, CD133+ and CD133+high.

254x190mm (200 x 200 DPI)